EXHIBIT K

1 COOLEY LLP
   BOBBY GHAJAR (198719)
2 (bghajar@cooley.com)
   TERESA MICHAUD (296329)
3 (tmichaud@cooley.com)
   COLETTE GHAZARIAN (322235)
4 (cghazarian@cooley.com)
   1333 2nd Street, Suite 400
5 Santa Monica, California 90401
   Telephone: (310) 883-6400
6
   MARK WEINSTEIN (193043)
7 (mweinstein@cooley.com)
   KATHLEEN HARTNETT (314267)
8 (khartnett@cooley.com)
   JUDD LAUTER (290945)
9 (jlauter@cooley.com)
   ELIZABETH L. STAMESHKIN (260865)
10 (lstameshkin@cooley.com)
   3175 Hanover Street
11 Palo Alto, CA 94304-1130
   Telephone: (650) 843-5000
12
   CLEARY GOTTLIEB STEEN & HAMILTON LLP
13 ANGELA L. DUNNING (212047)
   (adunning@cgsh.com)
14 1841 Page Mill Road, Suite 250
   Palo Alto, CA 94304
15 Telephone: (650) 815-4131

16 *[Full Listing on Signature Page]*

   Counsel for Defendant Meta Platforms, Inc.
17

18   **UNITED STATES DISTRICT COURT**

19   **NORTHERN DISTRICT OF CALIFORNIA**

20   **SAN FRANCISCO DIVISION**

21 RICHARD KADREY, *et al.*, | Case No. 3:23-cv-03417-VC-TSH

22   Individual and Representative Plaintiffs, | **DECLARATION OF PROFESSOR MICHAEL
                                                 | SINKINSON, PH.D. IN SUPPORT OF META'S
23   v. | MOTION FOR SUMMARY JUDGMENT**

24 META PLATFORMS, INC., a Delaware
   corporation;
25
                        Defendant.
26

27

28

I, Michael Sinkinson, Ph.D., declare:

1.   I am over the age of 18 and am competent to make this declaration. I have been engaged by Meta Platforms, Inc. ("Meta") as an economist to provide my opinion regarding certain economic factors relevant to fair use, in particular, (i) whether Meta's use of datasets containing Plaintiffs' books to train Meta's Llama models had any effect on the market for those books, (ii) whether a market for licensing Plaintiffs' books as training data exists or is likely to exist, and (iii) the societal benefits of Meta's Llama models.

2.   I submitted an opening expert report in this case on January 10, 2025 and a rebuttal report on February 3, 2025, and was deposed in this matter on March 4, 2025. The opinions below provide a more concise subset of the opinions set forth in my two expert reports. For convenience and ease of reference, I have included footnotes in each paragraph or section indicating where the text appears in my Opening and/or Rebuttal reports. The statements in my Opening and Rebuttal Reports are true and accurate to the best of my knowledge, information, and belief. Except as otherwise stated herein, I make this declaration based on my personal knowledge and professional expertise, Meta documents (as disclosed in my reports), the January 10, 2025 expert report of Professor Lyle Ungar, and research and experiments conducted by me or under my direction and supervision.

Professional Background and Qualifications

3.   I am an Associate Professor of Strategy at Northwestern University's Kellogg School of Management specializing in Applied Microeconomics and Industrial Organization and I am a Research Associate at the National Bureau of Economic Research (NBER). My previous appointments include the position of Associate Professor of Economics at the Yale School of Management, Assistant Professor of Business Economics and Public Policy at The Wharton School at the University of Pennsylvania, and as an Impact Fellow at the Federation of American Scientists.[1]

---

[1] Sinkinson Opening Report, ¶ 5.

1   4.   I previously served as an advisor to the Executive Office of the President of the

2   United States as Senior Economist at the Council of Economic Advisers (CEA) from July 2022 –

3   July 2023, where I was involved in competition and technology policy, which included representing

4   the CEA on the National AI Task Force. My contributions included shaping AI policy for federal

5   agencies and government contractors. The task force developed policies governing AI tool usage

6   within federal agencies and established rules for federal contractors. My efforts resulted in

7   executive orders and internal policy documents addressing AI regulation, incorporating evaluations

8   of costs, benefits, and priorities in the AI industry. My work also addressed critical issues such as

9   national security risks, economic impacts, broader innovation effects, and U.S. competitiveness in

10   the global economy. My research focuses on studying the structure of markets, particularly those

11   in the media, technology, and telecommunications industries. My research covers a number of

12   economic topics, including advertising, contracting, product positioning, and spectrum allocation.

13   I have co-authored several articles published in leading journals in economics, including the

14   *American Economic Review, The Review of Economic Studies, American Economic Journal:*

15   *Microeconomics, American Economic Review Papers & Proceedings, Quantitative Marketing and*

16   *Economics,* and the *Journal of the European Economic Association.* My research has been covered

17   by various media outlets, including *The New York Times*, *Bloomberg*, *The Economist*, and *Harvard*

18   *Business Review*.[2]

19   5.   I hold a Ph.D. in Business Economics from Harvard University's Harvard Business

20   School, and a Master of Arts in Economics from Harvard University. I received my Bachelor of

21   Commerce from Queen's University, Kingston, Ontario.[3]

Overview of LLMs

23   6.   Large Language Models (LLMs) are a class of generative AI ("Gen AI") models

24   that are trained on datasets of text. Once the LLM is trained, it takes a text prompt as an input and

25   then generates output in the form of a text response to that prompt. LLMs use deep learning

[2] Sinkinson Opening Report, ¶ 6.

[3] Sinkinson Opening Report, ¶ 7.

DECLARATION OF PROF. SINKINSON
CASE NO. 3:23-CV-03417-VC-TSH

techniques to capture the relationship between characters, words, and sentences to produce comprehensible text responses.[4]

7.   LLMs are trained on vast amounts of data across different domains and formats to develop a broader and more nuanced understanding of human language, enabling them to respond in natural human language. An LLM that is trained for general-purpose text generation, rather than for a specific task, is called an LLM *foundation model*.[5]

8.   The quality of output generated by an LLM depends on a number of factors, including the content and quality of the training data, the complexity of the model architecture, and the available computational resources. Larger models and more diverse data generally lead to model better performance, but they also increase the computational cost of both the training and deployment of an LLM. As LLMs continue to grow in size and capability, finding efficient ways to train and deploy them will be critical. LLM training data is typically collected from diverse sources, including websites, articles, research papers, books, and other collections of text. This diversity is key to making the models versatile, enabling them to handle different languages, topics, and tasks. And this text is often drawn from a wide variety of publicly-available sources. For example, one well-known LLM dataset known as "The Pile" includes a dataset known as the "Enron Emails," a corpus of hundreds of thousands of internal emails from Enron Corporation that were made public after a government investigation and the company's bankruptcy due to allegations of accounting fraud.[6]

9.   LLM technology is built on decades of research and development. LLM models have progressed from early statistical language models to the large-scale models used today. Although significant advances in deep learning technology happened more than 40 years ago, starting around 2010, recent increases in computational power due to complementary innovations in computer hardware have led to increased interest in neutral network statistical models. The latest LLM technology was enabled by transformer model architecture, which was first described by

---

[4] Sinkinson Opening Report, ¶¶ 19-20.

[5] Sinkinson Opening Report, ¶ 26.

[6] Sinkinson Opening Report, ¶ 29.

researchers at Google in 2017. These created an environment where vast amounts of data, compute (or computing power) and advanced models could be combined to lead to high quality general purpose LLMs.[7]

10.   The November 2022 launch of ChatGPT by OpenAI, a publicly accessible chatbot powered by GPT-3.5, put LLMs in the public spotlight as it gave people across all levels of expertise the ability to rapidly receive human-like text responses to a wide variety of text prompts. Meta has been releasing a sequence of models called Llama since early 2023. Google released its PaLM model publicly in March 2023, followed by releases of its Gemini model in late 2023 and early 2024. Other notable developers of LLM foundation models include startups Anthropic and Cohere.[8]

11.   The rapid increase in the complexity of LLM models—as illustrated by the numbers of parameters used by LLM models—illustrates the rapid advancements being made in this area of computer science. For example, Meta's recent Llama releases include both a relatively small model with 1 billion parameters and a model larger than all other previous releases at 405 billion parameters. Llama's small-parameter model allows for less computationally intensive model development, for example, for software that can be run on mobile devices. The larger models are meant to maximize performance and are suitable for use cases where output quality is of paramount importance. This parallel approach to LLM development illustrates the rapidly changing LLM landscape, as well as uncertainty about the range of future use cases for LLM models.[9]

<u>**There is No Evidence That Meta's Release of the Llama
Models Affected Sales of Plaintiffs' Books**</u>

12.   The original market for Plaintiffs' books comprises the buying, selling, and renting of such works in various formats such as print, eBooks, and audio. As I will explain below, in my opinion, sales of books on which Llama was trained, including Plaintiffs' books, have not been negatively affected by Llama's release. In economics, a set of goods (or services) can be considered

---

[7] Sinkinson Opening Report, ¶ 31.
[8] Sinkinson Opening Report, ¶ 32.
[9] Sinkinson Opening Report, ¶ 33.

substitutes if they can satisfy the same or a similar need. From an economic standpoint, evaluating substitutability can be done by considering whether the secondary (allegedly infringing) use reduces the demand for the original work because the same protected expression can now be obtained through multiple sources. This can reduce the creator's ability to monetize the work, reducing incentives to create new works. A lack of (actual or potential) substitution between the original work and the secondary use, on the other hand, indicates that there is no harm to the creator of the original work.[10]

13.   Based on the works that Plaintiffs allege were used to train the Llama models, in my opinion, the relevant market for understanding substitution is the market for the consumption of *trade books*, a term in the book publishing industry that generally refers to books targeted toward a general audience and available through a variety of sales channels.[11] Table 1, below, provides a summary of the genres of Plaintiffs' books whose copyrights were allegedly infringed by Meta.

Table 1. Summary of Genres Among Plaintiffs' Books Alleged in Complaint

	Genre [A]	Count [B]
[1]	Fiction	41
[2]	Memoir	3
[3]	Lifestyle & Practical Guides	3
[4]	Social Issues	2
[5]	**Total**	49

14.   As shown above, the majority of these works are fictional works, with a few falling into other categories, such as lifestyle books or memoirs, but all incorporate elements of creative storytelling, and these books were intended to be sold to a broad and general audience.[12]

15.   One incentive for creating trade books is to sell them, which authors often do through publishers.[13] The U.S. book publishing industry is large, with approximately 3.3 million new international standard book numbers ("ISBNs") registered in 2022, and some estimates from

---

[10] Sinkinson Opening Report, ¶¶ 14, 74, 91.

[11] Sinkinson Opening Report, ¶ 76.

[12] Sinkinson Opening Report, ¶ 77.

[13] Sinkinson Opening Report, ¶ 80.

2010 have placed the number of distinct printed books at close to 130 million.[14] There are five major publishing houses, including Penguin Random House, HarperCollins, Macmillan, Simon & Schuster, and Hachette, which collectively account for roughly 80 percent of the U.S. market for trade books, while the remaining 20 percent of the market includes a variety of small-to-medium publishing houses.[15]

16.   Other pecuniary incentives for creating trade books arise from additional markets derived from the market for trade books. For example, authors may benefit from their works being adapted into movies, screenplays, and videogames and from the production of merchandise. There is no reason to expect that such additional markets would be negatively impacted by the alleged use of trade books for LLM training, as I do not find any empirical evidence for negative impacts on the original market for the sale of trade books from these transformative LLM training uses, as I discuss further below.[16]

17.   In economic terms, market substitution occurs when substantially the same content (i.e., the same story or facts, expressed in the same way) is consumed from a copy of the book rather than the original. This might occur, for example, when a reader substitutes an unauthorized near-verbatim copy for a hardcover original printing. Certain works—such as novels—are intended to be read from start to finish; reading a summary or excerpt from a novel is not a substitute for the experience of reading the original work.[17]

18.   The question of substitution is, in part, an empirical question that can be addressed using market data. In the context of the use of copyrighted works as AI training data, assessing substitution entails studying the effect of a secondary use on their original market. I performed just such an empirical investigation by directly measuring the effect of the public release of Llama on book sales. I found that the release of Llama—and thus, the alleged secondary use of Plaintiffs' books as AI training data—had no negative effect on sales of the original works.[18]

---

[14] Sinkinson Opening Report, ¶ 48.
[15] Sinkinson Opening Report, ¶ 51.
[16] Sinkinson Opening Report, ¶ 80.
[17] Sinkinson Opening Report, ¶ 81.
[18] Sinkinson Opening Report, ¶ 82.

19.   To estimate the potential effects of the release of Llama 3 on the sales of books allegedly used as training data for Llama, I relied on an econometric estimation methodology called first differences, which is commonly used in empirical settings to study changes in outcomes over time. My methodology is supported by robustness checks which demonstrate that the conclusions of my analysis do not depend on any particular set of assumptions.[19]

20.   My analysis focused on the effects of the release of Llama 3, rather than the earlier releases of Llama 1 on February 24, 2023 or Llama 2 on July 18, 2023. Meta integrated Llama into its products and services at scale for the first time with the release of Meta AI, an AI assistant based on Llama 3, on April 18, 2024. On the same day, Meta released a conversational AI interface on the website "Meta.ai" which is available to the general public independent of whether they use other Meta products or services. My understanding is that Meta did not integrate earlier Llama releases into its products and services at scale to create a general-purpose chatbot, nor did it offer access to Llama in a user interface comparable to "Meta.ai". In addition, the Llama 3 foundation model has seen the largest take up by the AI developer community. In contrast, Llama 1 had only limited availability; access to Llama 1 required an application process and was granted on a case-by-case basis to academic researchers, those affiliated with organizations in government, civil society, academia, and industry research laboratories. Similarly, the Llama 2 model was publicly available but did not see the same level of uptake by AI developers as Llama 3 and therefore provides less scope for a possible detrimental effect on demand for books. Consequently, if the alleged use of books as training data for Llama had an effect on the market for trade books, such effect was far more likely to arise following the release of Llama 3 than Llama 2 or Llama 1.[20]

21.   I also focused my analysis on the sales of Plaintiffs' books in the "Books3" dataset as I have been informed that the Plaintiffs allege that this dataset was used to train all of the Llama models (including Llama 3) and that it contains all of the Plaintiffs' at-issue books. I understand that these books are also currently available for purchase. My analysis consistently demonstrated that there was no detectable effect of the release of Llama 3 on the sales of books allegedly used to

---

[19] Sinkinson Opening Report, ¶ 96.
[20] Sinkinson Opening Report, ¶ 97.

1   train the Llama models. Therefore, I cannot associate the alleged use of books as AI training data

2   for Llama 3 with any detrimental impact on the sales of these books.[21]

3   22.   In my analysis, I studied book sales rank data obtained from Keepa, a data provider

4   that tracks the sales ranks of a wide range of Amazon products. Specifically, I studied the impact

5   of the release of Llama 3 on the Amazon sales rank of a large sample of books, including the

6   Plaintiffs' books allegedly used in training the Llama models. Although Keepa data focuses on

7   sales through Amazon, I found it pertinent to my analysis because Amazon remains the largest

8   online bookseller in the United States and the world. As the largest online retailer, sales though

9   Amazon represent a substantial percentage of book total sales, and thus provide a leading indicator

10   of broader industry sales trends. Book sales trends reflected on Amazon, in other words, should be

11   reflected across other major sales channels. For reasons that I have outlined above, I used the first

12   differences framework to estimate the impact of the Llama 3 release. My results, which are

13   precisely estimated and validated by several robustness tests, are consistent: I did not detect an

14   impact of Llama 3 on the sales ranks of books allegedly used to train these models and cannot

15   associate the release of Llama models with a detrimental impact on book sales.[22]

16   23.   Sales rank, by construction, describes a book's sales relative to other books. The

17   higher the sales rank number, the worse it is selling relative to other books; thus, a book ranked No.

18   1 is selling better than all other books for which data is available. A rise in the sales ranking of

19   treated books—in the present case, the books that were allegedly used as training data for the Llama

20   models—implies that some other books have relatively fallen in sales ranking, and vice versa. This

21   means that, in the analysis of sales rank data, these other books implicitly act as the control group,

22   even if data for this group is not available.[23]

23   24.   An expected feature of book sales is that they often exhibit a common lifepath: there

24   is a sharp rise in sales around the time of publication followed by diminishing sales over time. As

25   such, the typical book experiences a greater share of its total lifetime sales revenue in the early part

26

27   [21] Sinkinson Opening Report, ¶¶ 97-98.
     [22] Sinkinson Opening Report, ¶ 102.

28   [23] Sinkinson Opening Report, ¶ 107.

1   of its sales lifepath. Accordingly, sales ranks also exhibit this pattern: the typical book will be

2   ranked worse (i.e. a higher sales rank numerical value) over time as other titles are released and

3   overtake it in the rankings. Thus, the common sales lifepath presents an empirical obstacle in

4   estimating the impact of Llama's release because the typical book experiences the trend of slowly

5   falling in the rankings shortly after release. For this reason, most books will be ranked lower in the

6   period after Llama's release relative to the period before simply due to the passage of time,

7   irrespective of Llama's release. The challenge for an analysis based on sales ranks is to tease out

8   these common trends that are unrelated to the release of Llama.[24]

9   25.   I address this challenge by employing the first differences regression framework.

10   This framework examines the changes over time, rather than the levels, of the outcome of interest

11   (in this context, the sales rank of books). This means that, for each book, the researcher takes the

12   difference in sale ranks from one period to the next, and then estimates the model using that data.[25]

13   26.   When working with data on changes over time (like first differences), any

14   characteristics that are specific to each book and constant over time (such as the author, the genre,

15   or date of publication) are all accounted for: characteristics that are constant over time are

16   differenced out and do not affect the estimation results. In practice, to calculate the first differences

17   in this sales rank, I subtract from each week's sales rank the sales rank of the previous week, and

18   create a variable capturing the changes in ranks week-over-week for each book. In this way, I

19   overcome the challenge of a constant trend across books acting as a confounder.[26]

20   27.   The first-difference framework is a common regression framework that has been

21   widely used in economic research, including my 2011 paper "The Effect of Newspaper Entry and

22   Exit on Electoral Politics" which is published in the *American Economic Review*, one of the most

23   renowned economics journals. Based on the empirical question, contextual setting, and data

24

25

26

---

[24] Sinkinson Opening Report, ¶ 107.

[25] Sinkinson Opening Report, ¶ 108.

[26] Sinkinson Opening Report, ¶ 109.

1 constraints, I deem the first differences framework to be the appropriate framework to conduct this

2 analysis.[27]

3   28.   In my analysis, I am interested in detecting whether the release of Llama 3 caused,

4 in the period shortly following its release dates, a change in the typical sales rank pattern of books

5 which were allegedly used in training the Llama models. The focus on a shorter time interval

6 following the Llama 3 release allows my analysis to be less affected by naturally occurring industry

7 changes in sales rank affecting all books than it would be if I examined a longer time horizon.[28]

8   29.   I understand that the Plaintiffs have alleged infringement of 49 books, but in order

9 to determine the impact more broadly (if any) of the release of Llama 3 on book sales, my analysis

10 includes but goes beyond Plaintiffs' works. Meta's counsel made available to me metadata of

11 books used to train the Llama models from the "Books3" dataset. This data includes 139,587

12 entries which are marked as "books3". Out of these 139,587 entries of metadata, I was able to

13 obtain a sample of 84,948 books allegedly used in training the Llama models that are tracked by

14 Keepa in the period from January 2021 through November 2024. I removed those books where the

15 publication date was not provided and books with evidently incorrect values for sales rank. This

16 resulted in 79,821 books. I also kept those books published in 2010 or later, for a total of 58,601

17 books in the final sample for analysis, which includes 38 of the 49 books listed in the Plaintiffs'

18 Complaint. Since demand for the typical book is greatest at the beginning of its sales lifetime, this

19 early period was when there is greatest scope for a detrimental effect of Llama 3's release. Thus,

20 it is appropriate to focus on recently published books.[29]

21   30.   Applying a regression equation to the Keepa data, which I detail in my January 10,

22 2025 report, returns the following result of the first difference estimation for sales ranks of books

23 allegedly used in training the Llama models:[30]

24

25

---

26 [27] Sinkinson Opening Report, ¶ 110.

27 [28] Sinkinson Opening Report, ¶ 111.

    [29] Sinkinson Opening Report, ¶ 117.

28 [30] Sinkinson Opening Report, ¶ 117.

1
2
3
4
5
6
7
8
9
10
11
12
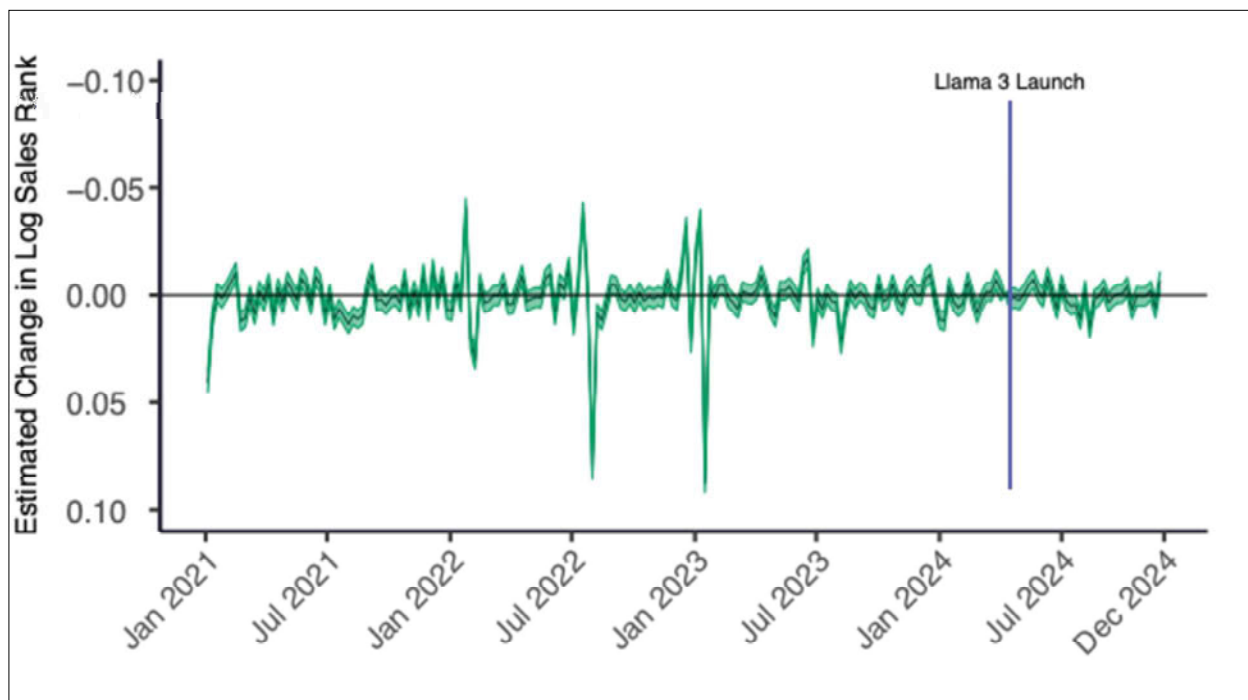
13   31.   The vertical axis has been inverted for ease of interpretation (negative coefficient

14 corresponds to better rank, and smaller numerical value of rank). The black line shows point

15 estimates from regression estimation. The green area is the 95% confidence interval. The point

16 estimate can be interpreted as the weekly change in log sales rank in a given week relative to the

17 weekly change in log sales rank in the week prior to Llama 3's launch.[31]

18   32.   As explained above, the release of Llama 3 is the most appropriate event on which

19 to focus my analysis. The figure above provides results from the first differences event study

20 estimated using a first differences regression equation.[32] It indicates a stable and flat first difference

21 pattern for the entire pre-release period, which supports the use of the first differences regression

22 framework for sales rank data. The graph also displays no visible changes in the weeks shortly

23 after Llama 3's release. The figure above does not display any outliers which would signify large

24 week-to-week changes in the weeks after the Llama 3 release. The coefficients of interest also

25 remain stable around zero after the Llama 3 release showing no permanent change in first-

26

27 [31] Sinkinson Opening Report, ¶ 117.

28 [32] Sinkinson Opening Report, ¶ 118. The first differences regression equation that was used is
    described in detail at Sinkinson Opening Report, ¶¶ 110-116.

difference pattern.[33]

33.   If the books under consideration had seen a sudden decrease in sales to a lower level and consequently a worsening in sales ranking, this would have been reflected in Figure 1 by a spike down shortly after the Llama 3 release. If there had been a consistent downward shift in the sales trend and consequently in sales rank, this would have been reflected by an estimated change consistently below zero after the Llama 3 release. The fact that these patterns are missing means I can conclude that book sales rankings continued on the same trajectory that they were on before the release of Llama 3.[34]

34.   I understand that the findings of my empirical analysis are consistent with statements from Plaintiffs. For example, I have reviewed transcripts of depositions of Plaintiffs Junot Diaz, Jacqueline Woodson, Rachel Louise Snyder, Sarah Silverman, Matthew Klam and others who testified that they are not aware of any decrease in sales of their books or other financial harm that can be attributed to the release of the Llama models. I am not aware of a single author who knows of any such decrease in sales of their books that can be attributed to the release of the Llama models.[35]

35.   In addition, counsel for Meta made available to me data on book sales and book royalties for some of Plaintiffs' works, which I understand was produced in discovery by Plaintiffs or their publishers. In reviewing this data, I observed patterns consistent with my analysis of Keepa sales rank data, such as the trends in average sales volumes qualitatively matching the average sales rank "lifepath" discussed further below. The data on Plaintiffs' works produced in discovery, however, differs substantially across Plaintiffs. For example, some Plaintiffs have provided data on sales volumes of their works, and other Plaintiffs have provided data on royalty payments. Some Plaintiffs have data over time, while others have produced only "snapshots" of data at one point in time. In addition to these inconsistencies within Plaintiffs' produced data, these data also contain information on only a small sample of works. Given their inconsistent quality and small size, the

---

[33] Sinkinson Opening Report, ¶ 118.

[34] Sinkinson Opening Report, ¶ 118.

[35] Sinkinson Opening Report, ¶ 103.

1   Plaintiffs' data does not allow me to draw an econometrically robust conclusion regarding potential

2   substitution behavior in the original market. Nevertheless, I did not find anything in reviewing that

3   data that was inconsistent with my conclusions derived from my regression analysis.[36]

LLMs Can Help to Promote Book Discovery and Improve Perceived Quality

5   36.   Books are in a category of goods that economists refer to as *experience goods*, which

6   is a term used for goods for which quality can be adequately evaluated only after the good has been

7   consumed. Whether a reader enjoys consuming a book largely depends on their particular tastes in

8   books. A consumer's book consumption experience depends on how they value features of a book,

9   such as the story, creativity, writing style and other factors. Both the relative importance of these

10   features, and a consumer's beliefs about how well a book delivers on these features, varies

11   substantially between individuals. These factors are not easy for a consumer to assess without

12   having read the book itself. In contrast, a consumer can assess the quality of a good *ex ante* when

13   a good is not an experience good. As an example, a buyer purchasing a new smartphone can

14   research the technical specifications of devices in the market, such as battery life, screen size, and

15   processing power, prior to making their purchase decision. The technical specifications of one

16   smartphone can easily be compared to those of other smartphones, allowing for a more fully

17   informed purchasing decision than a consumer might make when purchasing experience goods.[37]

18   37.   Book customers may still seek to inform their purchase first to ensure they make a

19   purchase which they believe will have high perceived quality. A book has high perceived quality if

20   it aligns with the consumer's idiosyncratic interests, preferences, and needs. Objective information

21   about books such as genre, number of pages, and reading level are all factors that can inform a

22   purchasing decision. A reader who is interested in a crime novel can restrict their attention to this

23   genre. A reader can also consider the title and back cover text, but these are subjective and provide

24   limited information. The nuances of writing style and story quality are difficult to infer and

25   therefore to access prior to the consumer making their purchase decision.[38]

---

[36] Sinkinson Opening Report, ¶ 103.

[37] Sinkinson Opening Report, ¶ 134.

[38] Sinkinson Opening Report, ¶ 135.

38.   Considering the difficulty to assess perceived quality *ex ante*, a range of options such as best-seller lists and expert reviews have emerged to help customers make more informed decisions. More generally, the digital economy offers a range of tools such as reviews and ratings to facilitate the consumption choice as I discuss further in this section. The major book retailers such as Amazon and Barnes & Noble provide this information on their product specific webpages. Other third-party websites such as Goodreads have also emerged to help consumers make informed decisions about their purchase through reviews and recommendation tools.[39]

39.   These options to acquire information about a book highlight the value of information transmission in the market between book readers and sellers. High-quality information on books helps customers predict whether they will enjoy reading a book, effectively reducing uncertainty about perceived quality. Any reduction of uncertainty on the buyer side can increase the demand for books.[40]

40.   For a consumer to inform their purchase with available information, they must first discover the book. There are millions of books that a reader could consider reading, but consumers are aware of only a small subset of those books. Consequently, discovering books that could be a good match for the consumer's interests is a prerequisite for purchasing books that will lead to a positive consumer experience.[41]

41.   Economists have found empirical evidence that an increase in perceived quality of books can lead to an increase in demand for books, resulting in an improvement for both book consumers and sellers. Economic theory suggests that demand increases when perceived quality increases, which puts upward pressure on prices. At the same time, high quality information about books increases competition among book sellers. An empirical investigation of online books markets shows that the positive effect on perceived quality outweighs the increase in competition such that demand and prices for books increase. As a result, consumers have better consumption

---

[39] Sinkinson Opening Report, ¶ 136.
[40] Sinkinson Opening Report, ¶ 137.
[41] Sinkinson Opening Report, ¶ 138.

1   experiences and sellers are also better off.[42]

2   42.   Tools that help readers search for new books or learn more about books that they

3   are considering reading help them in the purchase process. Such tools are an important part of

4   identifying goods to consume when a consumer is considering purchasing experience goods. In the

5   case of books, these tools allow the consumer to gauge the writing style, plot, content, and other

6   aspects that affect the consumer's perceived quality of a book to ensure that it will be enjoyable or

7   satisfying to consumers. By providing information about a book, these tools can help the user make

8   a more informed and confident purchasing decision.[43]

9   43.   Economists have documented evidence that the ability to search for text within books

10   increases sales of such books relative to non-searchable books. When readers can browse the full

11   text of the book for words or phrases but only view "snippets" (i.e., short excerpts), the evidence

12   shows that full-text digitization and searchability improves sales ranking. Searchability within a

13   book can also promote the discovery of authors and their other works.[44]

14   44.   Book snippets and excerpts are frequently included in professional book reviews,

15   such as those in the New York Times Book Review series, to help consumers assess whether a book

16   aligns with their preferences. The New York Times Book Review, for example, often features

17   carefully selected snippets from the book, alongside options to engage further, such as listening to

18   audio recordings of passages (e.g., the first chapter). The primary purpose of a book review is to

19   provide readers with insights to determine whether they wish to explore the book in its entirety,

20   rather than serving as a substitute for the book itself.[45]

21   45.   Advertisements in the movie industry and song samples in the music industry serve

22   as other examples for the value of excerpts to promote product discovery and consumption. The

23   wide use of movie trailers as advertisement for new releases indicates their positive impacts on box

24   office sales. Producers invest significant capital in commercial advertisements and carefully

25

---

26   [42] Sinkinson Opening Report, ¶ 139.

27   [43] Sinkinson Opening Report, ¶ 141.

     [44] Sinkinson Opening Report, ¶¶ 142-143.

28   [45] Sinkinson Opening Report, ¶ 144.

DECLARATION OF PROF. SINKINSON
CASE NO. 3:23-CV-03417-VC-TSH

develop trailers to draw viewer attention, interest, and curiosity in their new feature. At the same time, in the way that trailers are not substitutes for movies, excerpts and "snippets" (such as those that may be featured in book reviews or summaries) are not substitutes for reading books. In the music industry, access to song samples can help consumers to find a better match to their preferences when there is a great variety of songs to choose from.[46]

46.   LLMs can provide the same benefits to book authors and publishers that movie trailers do, by providing small samples of an artistic work, facilitating discovery, and providing more information about a product to potential consumers. By relying on online sources, an LLM may be able to provide a short passage of a book, raising awareness about the book.[47]

47.   In addition, Gen AI can enhance search technology to provide more accurate, contextually relevant search results and generate content that aligns closely with user queries. Google, the leading entity in general search services, uses Gen AI in their search functionality. LLMs can therefore be used as a technology to search queries within and across books when a customer is searching for books to purchase.[48]

48.   While LLMs may be able to provide search functionality or provide excerpts of books in which a consumer may be interested, the user cannot consume the full text of a book directly by prompting an LLM. As such, enhanced searchability and summarization does not result in a substitute for book consumption.[49]

49.   In today's digital economy, reviews, ratings, and book reception play a crucial role in shaping consumer behavior. They serve as a form of social proof, helping potential customers make informed decisions based on the experiences of others. Reviews and ratings rely on what economists refer to as the principles of *social learning* whereby individuals observe and imitate the behaviors and decisions of others to make informed choices. Potential consumers of experience goods can benefit from reviews and ratings to learn about whether they believe that a particular

---

[46] Sinkinson Opening Report, ¶ 145.

[47] Sinkinson Opening Report, ¶ 146.

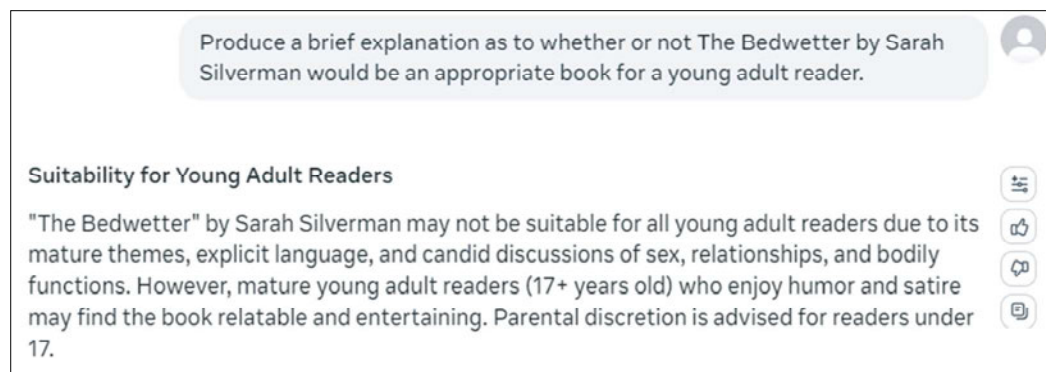[48] Sinkinson Opening Report, ¶ 147.

[49] Sinkinson Opening Report, ¶ 148.

1 good fits their particular interests.[50]

2   50.   LLMs have been shown to perform well at summarizing text according to human

3 evaluators. LLMs can produce new, personalized reviews to inform consumers, infer patterns from

4 text such as target audience demographics, reading level, or literary themes a reader may be

5 interested in. Because LLMs target their response directly to the user input, they have the capability

6 to provide reviews which include information that is specifically relevant to the user. For example,

7 when prompted generally for a review of the book *The Bedwetter* by Sarah Silverman, Meta AI

8 (which I am informed is powered by Llama) produces the response shown below.[51]

9

10

11

12

13

14

15

16   51.   The user could also have provided more detailed or nuanced prompts, asking for

17 information about the book from a particular perspective or viewpoint. For example, when

18 prompted about whether that same book would be suitable for young adult audiences, Meta AI

19 produced the output shown below.[52]

20

21

22

23

24

25

26

---

[50] Sinkinson Opening Report, ¶ 149.

[51] Sinkinson Opening Report, ¶ 155.

[52] Sinkinson Opening Report, ¶ 156.

DECLARATION OF PROF. SINKINSON
CASE NO. 3:23-CV-03417-VC-TSH

52.   AI can also promote book discovery and increase perceived quality by recommending books to users. Recommender systems help customers discover new products and are an integral part of ecommerce, including the online books market. Economists have found effects in the Amazon book network of recommendations, observing that the "co-purchased" information on a page leads to positive effects on book sales rank. Recommender systems are complex systems that rely on algorithmic understanding of customer's preferences. Gen AI as an emerging technology can help recommend new books based on its association to a book that a user previously read, a prompt, or any other data available.[53]

53.   In the example shown below, for example, Meta AI recommended a number of books in response to a prompt, including books by the Plaintiff author for which the user indicated that they liked one of their books, by recommending books from authors with a similar style.[54]
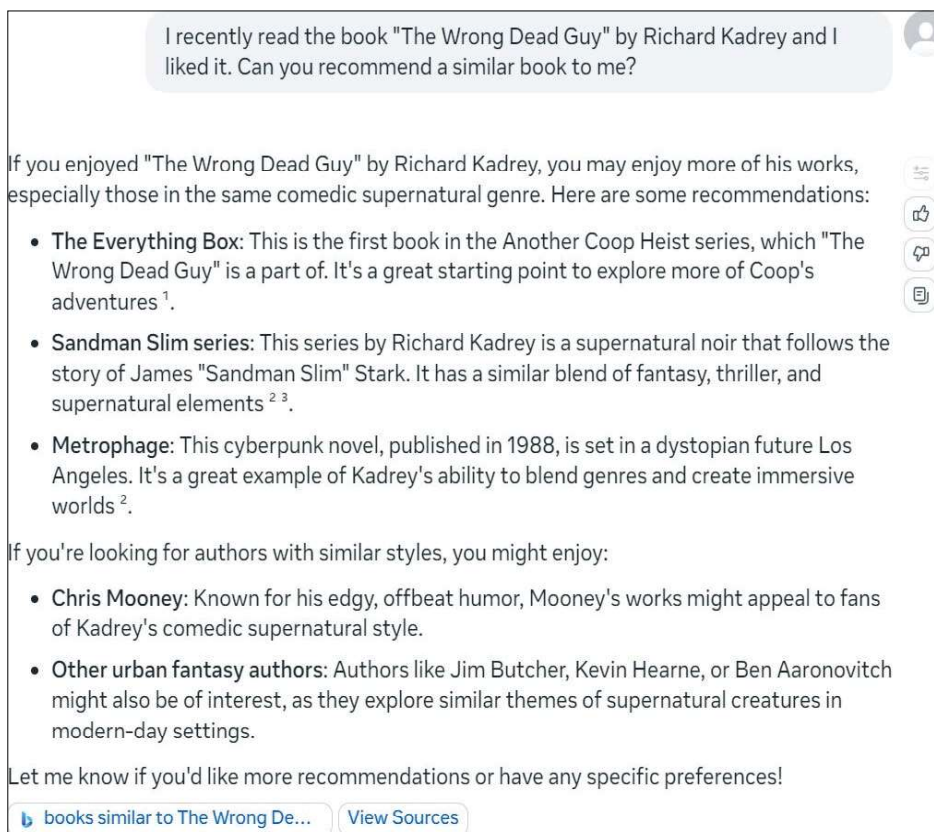
54.   The uniqueness of LLM responses can provide a novel way of book discovery and therefore increase the value of books in the original book sales market. As the evidence on the

---

[53] Sinkinson Opening Report, ¶ 158.

[54] Sinkinson Opening Report, ¶ 158.

DECLARATION OF PROF. SINKINSON
CASE NO. 3:23-CV-03417-VC-TSH

increase of book sales with differing ways of discovery suggests, LLMs can help users inform their purchasing decision and ultimately lead to an increase in book sales. This increase in demand for book purchases means that the book's value in its original market can potentially rise with the release of LLMs.[55]

A Hypothetical AI Licensing Market Is Likely To Fail

55.   In economics, a set of similar transactions between buyers and sellers is referred to as a *market*. *Market failure* is the economic term used to describe the conditions under which a set of potential buyers value a good and a set of potential sellers would be willing to sell that good, but either no transactions occur, or the quantity of goods transacted is *socially inefficient*. There are several reasons that market failure can happen.[56] One reason that market failure occurs is due to the presence of transaction costs that are too high.[57] Another complementary reason market failure can occur is when the value of a transaction is uncertain, making it difficult for buyers and sellers to reach agreements to trade.[58] There are at least two possible ways to view a hypothetical AI licensing market: (1) as focused specifically on licensing of the Plaintiffs' books (and other trade books)[59] for LLM training or (2) expanded to include licensing data generally for LLM training.[60] Under either formulation, as explained below, these hypothetical markets are likely to fail.

56.   ***First***, the value of data to LLM developers in training LLM products varies based on the type of information it contains. While the value of books to LLM developers is likely to differ based on language, genre, topic, author, length, and other features of the work, the marginal value of any one work—no matter what combination of features it has—is very low. LLMs are trained on thousands (or even millions) of books, with *trillions* of individual pieces of data used to calibrate modern LLMs. This volume of data is not a "nice to have," but a "must" for training

---

[55] Sinkinson Opening Report, ¶ 160.
[56] Sinkinson Opening Report, ¶ 181.
[57] Sinkinson Opening Report, ¶ 182.
[58] Sinkinson Opening Report, ¶ 183.
[59] Sinkinson Opening Report, ¶ 180.
[60] Sinkinson Rebuttal Report, ¶¶ 90-91.

LLMs, since LLMs become usable only when trained on massive amounts of data.[61] I understand from Professor Ungar that the immense size of LLM training datasets "is crucial because LLMs begin with random weights and must learn everything about language from scratch—word meanings, syntax, structure, and world knowledge—from statistical patterns in text alone. Achieving this requires a massive amount of diverse data to capture the complexity and variety of human language, and to be able to effectively generalize what has been learned through training."[62] For example, Llama 3.1 has three major versions, each of which corresponds with a different number of model parameters that dictate the output of the model. The smallest of these models was trained on 15 trillion tokens (or roughly 11.25 trillion words). In order for LLMs to understand the relationships between words in a sentence and recognize the broader stylistic and grammatical patterns in written language, they require vast volumes of training data.[63]

57.   I also understand from Professor Ungar that given the enormous volume of text necessary to train LLMs, one additional book does not significantly improve an LLM, and therefore the marginal value of one additional book (or a handful of books) from the perspective of the LLM developer is also vanishingly small. His analysis and results demonstrate that the inclusion of any individual book has a negligible, if any, effect on the overall capabilities of the LLM. From an economic perspective, if an individual book does not improve model performance, then it has little, if any, economic value as training data. From a market standpoint, the negligible value of individual books significantly diminishes the incentive and willingness of LLM developers to engage in the expensive and time-consuming process of engaging in bilateral negotiations with numerous third parties, such as individual authors and others, who may at most only be able to license one or perhaps a handful of books for LLM training.[64]

58.   *Second*, transaction costs associated with licensing trade books as AI training data are likely to be high because they require negotiations with a large number of individual authors in

---

[61] Sinkinson Opening Report, ¶¶ 193-194.

[62] Sinkinson Rebuttal Report, ¶ 91 (quoting Ungar Opening Report, ¶ 91).

[63] Sinkinson Opening Report, ¶ 194.

[64] Sinkinson Opening Report, ¶ 198.

order to obtain a sufficient volume of books to train an LLM. As noted previously, there are five major publishing houses, including Penguin Random House, HarperCollins, Macmillan, Simon & Schuster, and Hachette, which collectively account for roughly 80 percent of the U.S. market for trade books, while the remaining 20 percent of the market includes a variety of small-to-medium publishing houses.[65] In the book publishing industry, the term *subsidiary rights* describes the rights beyond the primary book publishing rights. These include the rights to license or produce translations, audiobooks, film and TV adaptations, merchandise, and more. When negotiating publishing agreements, authors may choose to retain these rights for themselves so they can negotiate separate deals, potentially increasing and diversifying the revenue streams from their works.[66]

59.   Critically, subsidiary rights for licensing works to train LLM models are not a clearly delineated right in existing publishing contracts. I understand that almost all of the Plaintiffs in this case testified at deposition that they understood their publishing agreements to reserve to them, not their publisher, the right to license their works for AI training.[67]

60.   The dispersion of subsidiary rights for licensing will further increase the transaction costs of a bilateral negotiation between the LLM and the author/publisher. First, the LLM developer has no way of knowing the party with whom it would need to negotiate the rights to use a literary work as LLM training data. Second, the author and publisher themselves may not even agree over who possesses this right due to ambiguity in the publishing agreement. Third, it is speculative as to whether any initiatives by publishers to negotiate on behalf of authors will significantly reduce transaction costs such that the mass scale of licensing LLM training data will be feasible, as doing so would depend not only on LLM developers' ability to get consent from authors, but to get consent for a sufficiently large number of works (and the right mix of works) to

---

[65] Sinkinson Opening Report, ¶ 51.

[66] Sinkinson Opening Report, ¶ 202.

[67] Sinkinson Opening Report, ¶ 203. For example, none of the publishers of Plaintiffs' works is a plaintiff in this case and I am unaware of any publisher claiming to have ownership over the rights to license the 13 Plaintiffs' works for LLM training.

make this licensing activity worthwhile to LLM developers.[68]

61.   Given the variation in rights granted to publishers versus authors across publishing agreements, as well as the other complexities of the book publishing market, LLM developers would have to invest significant time and resources to undertake each unique licensing negotiation. The cost of determining which party holds the licensor rights for LLM training, and successfully negotiating licenses for these rights, would very likely exceed the minimal value each licensed work provides to the LLM developer. Such an outcome would result in market failure.[69]

62.   Plaintiffs' economic expert, Prof. Daniel Spulber, discusses a range of collective licensing organizations in his January 10, 2025 opening report, which, in his opinion, is evidence that there is a broad market for licensing data. However, all of those organizations cover uses separate from AI training data or are otherwise irrelevant to the question of whether or not a market that includes licensing books for use as AI training data exists.[70] Despite an expansive review of collective licensing organizations, Prof. Spulber does not present a single collective licensing agreement that would enable licensing the Plaintiffs' works or trade books more generally as AI training data.[71]

63.   Prof. Spulber's broad market definition also does not distinguish between the *service* of providing access to data, and licensing individual books. Many of the agreements cited by Prof. Spulber as support that there is a market for Plaintiffs' books concern the provision of access to large volumes of data, with several of them including explicit acknowledgements that such access does not constitute an intellectual property license.[72] For example, Dr. Spulber observed that in October 2023, ███████████████████████████████

██

███. As these materials are not copyrighted, any fees to be paid under this agreement are unrelated to a copyright license, which

---

[68] Sinkinson Opening Report, ¶ 204.

[69] Sinkinson Opening Report, ¶ 209.

[70] Sinkinson Rebuttal Report, ¶¶ 81-87.

[71] Sinkinson Rebuttal Report, ¶ 81.

[72] Sinkinson Rebuttal Report, ¶ 15.

underscores the value of data access independent of any licensing of intellectual property.[73] While this may be evidence that there is a market for providing access to large, curated databases of text, it does not indicate that there is a market to license individual books.[74]

64.   ***Third***, uncertainty about the future of LLMs increases transaction costs. In a hypothetical market for LLM training data, different trade books would have wide variation in buyer willingness to pay and seller willingness to receive due to the differing circumstances of both the LLM developer and the author/publisher of trade books. LLM development is still in the early stages, with LLM market entry occurring rapidly and LLM developers making major updates to their models on a regular basis. For example, Meta's Llama has been through more than three substantive iterations in just over 18 months. Llama 1 was released in February 2023, with models that used between 7 and 65 billion parameters. Only five months later, Meta released Llama 2 in July 2024, which used 70 billion parameters. Within a year, Meta released Llama 3, Llama 3.1, which drastically improved on previous models with a 405 billion parameter model, Llama 3.2, and Llama 3.3. During this period, many other companies have also entered the Gen AI LLM market. This rapid growth in the number of market participants, as well as the advancement in model size and capability in just over a year, demonstrates the dynamic nature of the LLM market.[75]

65.   In addition to the dynamic state of LLM development, the LLM market has not sufficiently matured such that developers understand what use cases will prove most profitable in the long run, the methods of monetization that will prove the most viable, and what types of training data will be used to train models that users will most demand. For example, as LLMs mature, the demand for LLMs that are trained mainly on scientific documents or up-to-date news content may be more in-demand than LLMs that are trained on books. In the same way that LLM developers' willingness to pay for hypothetical data licenses is likely to shift as the LLM market matures, authors/publishers are also likely to change their willingness to receive for such licenses. Authors and their publishers are likely to update their beliefs about the economic value of their works being

---

[73] Sinkinson Rebuttal Report, ¶¶ 15, 39.

[74] Sinkinson Rebuttal Report, ¶ 15.

[75] Sinkinson Opening Report, ¶ 210.

1   used in AI training data.[76]

2   66.   Given these market dynamics, both buyer willingness to pay and seller willingness

3   to receive in a hypothetical market for AI training data licenses is likely to shift in the near future.

4   This increases transaction costs in multiple ways. First, these market dynamics will cause both

5   parties to be less willing to agree to be locked-in to long-term contracts in the event that they later

6   want to renegotiate any licensing agreements to more favorable terms in response to changes in

7   their willingness to pay or willingness to receive. Shorter term contracts lead to more frequent

8   bilateral negotiations, which in turn increase transaction costs.[77]

9   67.   I have reviewed the transcripts (with exhibits) of Meta employees Alex Boesenberg,

10   Amanda Kallet, and Sy Choudhury, relating to efforts by Meta in early 2023 to discuss possible

11   licensing with book publishers, which I understand were ended in or about April 2023. The

12   responses that Meta received to these discussions are consistent with my opinions regarding the

13   lack of an AI licensing market for the Plaintiffs' works. In many cases, publishers with whom Meta

14   was communicating lacked rights to license entire catalogs for AI training. As a result, many

15   agreements would have required individual author consent.[78]

16   68.   As explained above, a market for licensing trade books as training data suffers from

17   market failure for a variety of reasons, including, but not limited to, dispersion of rights among a

18   voluminous number of rightsholders; uncertainty about the economic benefits of LLMs; the *de*

19   *minimis* value, if any, that any individual work contributes to model training; and the difficulty of

20   appraising that (zero or near-zero) value, all of which contribute to high transaction costs and

21   market failure.[79]

22   69.   Plaintiffs' economic expert, Prof. Spulber, describes the relevant licensing market

23   broadly as "licensing in the market for LLM training data," which he claims is well-established,

24   large, and rapidly growing. But this definition appears to encompass any kind of data, and even if

25

---

26   [76] Sinkinson Opening Report, ¶ 211.

27   [77] Sinkinson Opening Report, ¶ 213.

     [78] Sinkinson Opening Report, ¶¶ 214-215.

28   [79] Sinkinson Rebuttal Report, ¶ 91.

DECLARATION OF PROF. SINKINSON
CASE NO. 3:23-CV-03417-VC-TSH

1   limited to textual data, it is so broad that it collapses in upon itself and underscores the infeasibility

2   of a licensing regime for LLM training data. All of the considerations I identified above pertaining

3   to why a hypothetical market for licensing books for LLM training was likely to fail, would make

4   an expanded hypothetical licensing market even more likely to fail.[80]

5   70.   Books are only a small category of data used to train LLMs. As Prof. Ungar explains

6   in his opening report, as noted above, LLM training datasets are massive, with modern LLMs

7   trained on between one to more than 15 trillion tokens. Aside from a large number of tokens, an

8   LLM also, accordingly, requires an enormous diversity of tokens to respond to a nearly infinite

9   range of potential prompts—much more diversity than exists within books. Yet all of the factors

10   that undermine the viability of a market that includes licensing books to train LLMs are exacerbated

11   for other categories of text training data. For example, the corpus Common Crawl, which was the

12   largest single source of pretraining text for Llama, contains approximately 15 trillion tokens of

13   deduplicated data composed of billions of discrete works that may or may not be subject to

14   copyright protection. Determining which portions of Common Crawl are protected by copyright,

15   let alone who owns those rights, and negotiating and executing agreements with those rightsholders,

16   is impossible.[81]

17   71.   Another factor exacerbating the lack of an AI licensing market is the speed at which

18   competition in the generative AI field is progressing, including by entities outside the United States.

19   For example, beginning in December 2024, large language models were recently announced and

20   made available by DeepSeek and Alibaba (QWEN 2.5), both developed by engineers claiming that

21   their models achieve performance that meets – and in some cases exceeds – competing LLMs, and

22   it is highly likely that other competitors will emerge in the near future. It will be problematic for

23   the development of LLMs and the competitiveness of LLM creators (especially in the United

24   States) if they have to wait for the emergence of a hypothetical future licensing market for AI

25   training data that is unlikely to ever exist (whether based on books or data generally), and in the

26   meantime, engage in licensing negotiations with a potentially countless number of third parties over

27

_____

[80] Sinkinson Rebuttal Report, ¶¶ 90, 92.

28   [81] Sinkinson Rebuttal Report, ¶ 91.

a period of years to obtain licenses for the requisite number of tokens needed to train modern LLMs.[82]

Meta's Open Release of Llama Benefits the Public

72.　　Llama's open-source nature allows many developers, including hobbyist or startup developers with limited computational resources, to fine-tune the foundational model for their particular use cases. For instance, a STEM education platform has fine-tuned the foundation Llama model to support young STEM students' scientific learning. Zoom, the popular video meeting platform, has leveraged Llama 2 in developing its Zoom AI companion feature, which allows users to summarize meeting notes to boost productivity.[83]

73.　　For open-source AI models, model weights can be accessed, downloaded and modified by any user. AI model weights are the numerical parameters which define the relationship between a model's data inputs and determine how the model makes predictions. Thus far, the leading Gen AI models—known as *frontier models*—including OpenAI's GPT and Google DeepMind, have been closed, that is, users are unable to access and download the model weights.[84]

74.　　Open-source foundational AI models increase accessibility and transparency, driving innovation and competition across the economy.　Compared to closed models, open-source models significantly reduce the level of technical knowledge needed to create AI applications, increasing the ability of small businesses to implement AI-powered tools at a faster pace and with lower cost barriers while preserving their ability to control their own data. By fostering a collaborative environment for developers, open-source models also help improve the quality of AI algorithms; since thousands of developers are working in the same ecosystem, software bugs are resolved more efficiently and more researchers can contribute to improving the software.[85]

75.　　Open-source AI algorithms also improve AI security and safety, for example, by increasing the likelihood that harmful biases are discovered and promptly addressed. Additionally,

---

[82] Sinkinson Rebuttal Report, ¶ 93.

[83] Sinkinson Opening Report, ¶ 45.

[84] Sinkinson Opening Report, ¶ 174.

[85] Sinkinson Opening Report, ¶ 175.

1   due to their public nature, open-source models can be widely and rigorously tested by third parties

2   and regulatory agencies. Open-source models may also prove to use fewer computational resources

3   in model training. The computing power required for training a model is directly correlated with

4   the cost of running it and the emissions released in this process. Since deploying open-source

5   models that have already been developed and training them for specialized purposes requires fewer

6   computational resources than building specialized models from scratch, developers can save

7   significantly on computational and energy costs. The lower cost of running open-source models

8   also creates opportunities for varying pricing models to provide consumers with the most

9   competitive prices and maximize economic gains overall.[86]

10   76.   Meta has led the push for open-source AI and partnered with several government

11   agencies and other industry leaders to promote AI applications in the public sector. On July 23,

12   2024, Meta released Llama 3.1 405B, the first frontier-level open-source LLM. With this release,

13   Mark Zuckerberg released a letter stating "open source AI represents the world's best shot at

14   harnessing this technology to create the greatest economic opportunity and security for everyone."

15   Since the White House roundtable on U.S. leadership in AI infrastructure, Meta has also made its

16   models available to all government agencies and several government partners. Oracle, ScaleAI, and

17   Lockheed Martin are all using Meta's open-source model to develop national security and defense

18   programs, as well as to supplement existing data analysis and code generation functions.[87]

19   77.   The White House has repeatedly emphasized the importance of artificial intelligence

20   as part of bolstering America's competitive advantage in the technology sector. Successive

21   administrations have issued Executive Orders and memos aimed at promoting AI innovation and

22   warning against disadvantaging U.S. companies and the public by allowing geopolitical rivals to

23   pull ahead in AI development.[88]

24   78.   In October 2023, former President Biden issued an executive order (since rescinded

25   by the Trump administration) on the safe, secure, and trustworthy development of artificial

26

27   [86] Sinkinson Opening Report, ¶ 176.

       [87] Sinkinson Opening Report, ¶ 177.

28   [88] Sinkinson Opening Report, ¶ 165; see

intelligence. To ensure the United States' place as a continued global technology leader, the order highlighted the need to accelerate AI research, identify ways to incorporate AI safely and effectively, and promote open competition between technology firms. This was followed by an October 2024 National Security Memorandum titled "Advancing the United States' Leadership in Artificial Intelligence; Harnessing Artificial Intelligence to Fulfill National Security Objectives; and Fostering the Safety, Security, and Trustworthiness of Artificial Intelligence," which emphasizes the U.S. government's objectives of maintaining U.S. leadership in the development of advanced AI, accelerating the adoption of state of the art AI systems by national security agencies, and development of robust AI governance and risk management practices. Following the submission of my Opening Report, President Biden also issued a January 2025 executive order on "Advancing United States Leadership in Artificial Intelligence Infrastructure," which described the "imperative" of "building AI infrastructure in the United States on the time frame needed to ensure United States leadership over competitors who, already, are racing to take the lead in AI development and adoption."[89]

79.   I served on the National AI Task Force of the Biden presidential administration. In this role, I witnessed firsthand the guiding principle of regulating AI in a way that fosters innovation while ensuring America's technological leadership. I contributed to shaping AI policy for federal agencies and government contractors. The task force worked on policies governing AI policy and AI tool usage within federal agencies and on establishing rules for federal contractors—a significant lever given the economic influence of government contracting on a wide range of firms in the economy. The task force's efforts culminated in executive orders and internal policy documents addressing AI regulation, including evaluations of costs, benefits, and priorities. These analyses considered issues such as national security risks, economic impacts, broader innovation effects, and U.S. competitiveness in the global economy. During my tenure, I observed how AI

---

[89] Sinkinson Opening Report, ¶¶ 165-166; *see also* Exec. Order No. 14141, accessible at https://bidenwhitehouse.archives.gov/briefing-room/presidential-actions/2025/01/14/executive-order-on-advancing-united-states-leadership-in-artificial-intelligence-infrastructure/.

1 stands as a transformative innovation that brings significant value to the public, enhancing

2 industries, improving efficiencies, and addressing societal challenges.[90]

3   80.   In late January 2025, President Trump issued a new Executive Order titled

4 "Removing Barriers to American Leadership in Artificial Intelligence," which notes that the

5 "United States has long been at the forefront of artificial intelligence (AI) innovation, driven by the

6 strength of our free markets, world-class research institutions, and entrepreneurial spirit, and

7 declares it to be "the policy of the United States to sustain and enhance America's global AI

8 dominance in order to promote human flourishing, economic competitiveness, and national

9 security."[91]

10   81.   Attached hereto as Exhibit A is a true and correct copy of my current curriculum

11 vitae. It contains a complete list of publications, awards, and prior employment history.

12

13   I declare under penalty of perjury that the foregoing is true and correct. Executed on this

14 24th day of March, at Evanston, Illinois.

15   /s/
                                    Prof. Michael Sinkinson, Ph.D.
16

17

18

19

20

21

22

23

24

25

26

27 ──────────────
[90] Sinkinson Opening Report, ¶ 167.

28 [91] Exec. Order No. 14179, accessible at https://www.whitehouse.gov/presidential-
    actions/2025/01/removing-barriers-to-american-leadership-in-artificial-intelligence/.

COOLEY LLP
ATTORNEYS AT LAW

DECLARATION OF PROF. SINKINSON
CASE NO. 3:23-CV-03417-VC-TSH